Canadians Land on Jupyter

Dr. Ian Allison
Dr. James Colliander

Pacific Institute for the Mathematical Sciences
“The PIMS mandate is to promote excellent research and applications of the mathematical sciences, to facilitate the training of highly qualified personnel, to enrich public awareness of, and education in, mathematics, and to create partnerships with similar organizations around the world.”
How do we ensure that students and researchers have access to the tools they need?
syzygy

ˈsizijē/ noun ASTRONOMY; ˈsizijē/ noun MATHEMATICS; ˈsizijē/ noun BIOLOGY;

ˈsizijē/ noun; A collaboration between PIMS, Cybera and Compute Canada to make institutional JupyterHub’s available across Canada
syzygy: Design Goals

- Super simple access
- Automate and minimize admin
- Be hardware agnostic
- Partition resources
- Share everything!
Tools and Resources
Terraform - *Infrastructure as code*

Ansible - *Simple IT Automation*
syzygy - The Moving parts

- JupyterHub
- Shibboleth
- DockerSpawner
- ZFS, dehydrated, ...
Monitoring: Prometheus & Grafana
syzygy Today
By the Numbers

- 19 Hubs; 14 universities + “others”
- 3 clouds + some metal
- ~8000 users (400,000 have access)
- 5 TB of Disk used (/home + docker)
- 567 GB of memory
- 124 CPU cores
Testimonials

“Syzygy gives my students access to modern, cloud-based computing in a format that feels natural to them, even those unfamiliar with computing.”

“[Jupyter] matches the expectations and needs of students in many disciplines working with data to build arguments.”

- Meghan Allen & Steve Wolfman, CS103
Testimonials

“In 2017 and 2018, I co-hosted the BC Data Science Workshop, featuring an array of student teams collaborating with their respective industry mentors … Using syzygy and JupyterHub as the interface for the participants streamlined both their experience and ours. The configuration was painless and allowed us to minimize the number of moving parts with which students needed to be familiarized”

- Aaron Berk, PhD student, Organizer: BCdata workshop
Testimonials

“I estimate the number of hours students are engaged with software and data has increased 10-fold. Syzygy has removed significant barriers to entry for math students pursuing careers in data and computation.”

- Patrick Walls, S$^3$ + Math 210 Introduction to Mathematical Computing
syzygy Tomorrow
How many users can a Hub handle?

What about 10 * N, 100 * N, … ?

Plans - Technical

<table>
<thead>
<tr>
<th>Technical edge (nginx) (replicated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hubsharder (replicated)</td>
</tr>
<tr>
<td>$cookie</td>
</tr>
<tr>
<td>psql</td>
</tr>
<tr>
<td>Hub A1</td>
</tr>
<tr>
<td>Hub A2</td>
</tr>
<tr>
<td>…</td>
</tr>
<tr>
<td>Storage provider ZFS/NFS</td>
</tr>
<tr>
<td>Auth provider</td>
</tr>
</tbody>
</table>
● More syzygy hubs!

● Teaching and Research

● Social Collaboration Structures
  ○ Workshops
  ○ University-Industry
  ○ University-Government

Plans - Community
● Compute Canada Clusters

● github.com/cmd-ntrf/jupyter-lmod + CVMFS

● Clusters cloud-init

● Cloud-init deployments; Workshops
Effort > Expertise

- Most of this is implementation not development
- There’s a huge appetite for that implementation!

Canadians Have Landed on Jupyter.
Thank You